JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2006, **86,** 239–251 NUMBER 2 (SEPTEMBER)

TESTING RESPONSE-STIMULUS EQUIVALENCE RELATIONS USING DIFFERENTIAL RESPONSES AS A SAMPLE

Hirofumi Shimizu

MEISEI UNIVERSITY, TOKYO, JAPAN

This study tested the notion that an equivalence relation may include a response when differential responses are paired with stimuli presented during training. Eight normal adults learned three kinds of computer mouse movements as differential response topographies (R1, R2, and R3). Next, in matching-to-sample training, one of the response topographies was used to select a comparison stimulus B (B1, B2, or B3) conditionally upon presentation of sample stimulus A (A1, A2, or A3), and to select stimulus D (D1, D2, or D3) conditionally upon presentation of stimulus C (C1, C2, or C3). After two sample–comparison–response relations (ABR and CDR) were established, 18 sample–comparison relations were tested (BA, DC, RA, RB, RC, RD, AC, CA, AD, DA, BC, CB, BD, DB, AA, BB, CC, and DD). In the RA, RB, RC, and RD tests, the differential responses (R1, R2, and R3) were used as sample stimuli. All subjects made class-consistent comparison selections in the tests. This study provides evidence that responses may become members of an equivalence class.

Key words: response–stimulus relations, conditional discrimination, equivalence relation, differential responses, matching-to-sample, humans

---

Once a conditional relation between two or more stimuli is established by reinforcement, additional relations can emerge without explicit training. For example, following reinforcement selecting Comparison B conditionally upon Sample A (A–B relation), a subject may spontaneously select Comparison A conditionally upon Sample B (B–A relation). In equivalence relations, in which conditionally related stimuli share functional properties (e.g., Sidman, 1971), such emergent relations are especially plentiful. Much is known about the factors that mediate emergence of untrained stimulus–stimulus relations (e.g., Dube, Green, & Serna, 1993; Fields & Verhave, 1987; Sidman, Kirk, & Willson-Morris, 1985; Sidman & Tailby, 1982; Steele & Hayes, 1991), although stimulus-reinforcer relations also are known to emerge (e.g., Dube & McIlvane, 1995; Dube, McIlvane, Mackay, & Stoddard, 1987; Dube, McIlvane, Maguire, Mackay, & Stoddard, 1989; Goyos, 2000; McIlvane, Dube, Kledaras, de Rose, & Stoddard, 1992; Schenk, 1994).

Sidman (2000) proposed that equivalence relations may incorporate not only stimuli and reinforcers but also any differential responses that are paired with them. Figure 1 illustrates Sidman's argument. The upper portion of the figure shows six, four-term contingencies (Sidman, 1986) established by matching-to-sample training involving two sets of three samples, three comparisons and three responses. Selection of Comparison B1 (or B2 or B3) conditionally upon Sample A1 (or A2 or A3) with Response R1 (or R2 or R3) is reinforced (the same reinforcer is used in all cases). Additionally, selection of Comparison D1 (or D2 or D3) conditionally upon Sample C1 (or C2 or C3) with Response R1 (or R2 or R3) is reinforced. The lower left portion of Figure 1 shows the equivalence relations expected to result from this training, with emergent relations divided into three categories. The first category contains reflexive and symmetrical stimulus–stimulus relations, which are most typically the focus in equivalence experiments (e.g., Sidman & Tailby, 1982). The relations in the second category explicitly include responses, whereas those in the third category are stimulus–stimulus relations forged via association with a common response. The

The author served as a Research Fellow of the Japan Society for the Promotion of Science when this research was conducted. This research was supported by the Ministry of Education of Japan (Grant-in-Aid for JSPS Fellows H11-00148). The author thanks Professor Hirokazu Yoshimura of Hosei University for suggestions regarding the experimental design; Tina Covington, Program Supervisor at Hawthorne Country Day School, for support in finishing the manuscript; Janet S. Twyman for helpful comments on the manuscript; and William V. Dube for his total support. The author gratefully acknowledges the editorial process of this paper supported by Thomas S. Critchfield.

Requests for reprints should be sent to Hirofumi Shimizu (E-mail: hiro.shimizu@gmail.com).

doi: 10.1901/jeab.2006.04-03

latter two categories have been ignored in most equivalence experiments, which typically employ a single response topography in all phases of training and testing. In this example, note that 51 of 69 possible emergent relations incorporate, or are mediated by, differential response topographies.

Only a few published studies have implicated responses in equivalence class formation. Those conducted with humans typically have employed verbal responses. For instance, Eikeseth and Smith (1992) demonstrated the emergence of equivalence relations in preschool-aged autistic children after they were trained to apply a consistent label to members in the same class (for related findings, see Lowe, Horne, Harris, & Randle, 2002). Several studies with nonhumans, however, suggest that the common response need not be verbal. McIntire, Cleary, and Thompson (1987) used differential responses to demonstrate equivalence relations in macaques. The subjects were required to perform a differential response to a sample to produce comparisons, and they performed the same response when they selected a comparison. Consequently, they showed equivalence relations as defined by Sidman and Tailby (1982). Thus, stimuli appeared to become equivalent by virtue of association with a common response (for an alternative interpretation of this study, however, see Hayes, 1989).

Manabe, Kawashima, and Staddon (1995) trained budgerigars to make low- or high-frequency calls in response to different discriminative stimuli (C1 and C2); that is, the budgerigars learned to make a C1 call and a C2 call. After this training, a matching-to-sample task was introduced. Given sample S1 or S2, the budgerigars were required to make a call to produce comparisons (C1 and C2). Pecking on C1 (C2) conditionally on S1 (S2) was reinforced. Although the calls were not differentially reinforced, the budgerigars made the C1 call in the presence of S1 and the C2 call in the presence of S2, suggesting an equivalence relation that included a response (for an alternative interpretation of this study, see Saunders & Williams, 1998). Lionello-DeNolf and Urcuioli (2003) reported an analogous finding with pigeons, using patterns of key pecks as differential responses.

The present study, which was modeled after the hypothetical case in Figure 1 (see Sidman, 2000), was designed to explore the development of equivalence classes in humans involving a putatively nonverbal response. Differential responses were based on the manipulation of a computer mouse in a design that allowed most of the untrained conditional discriminations shown in Figure 1—including those in which a response served as sample—to be tested.

METHODS

Subjects

Eight subjects (ages 19 to 28 years), 4 females (MKB, SUT, TMM, and UIR) and 4 males (IOH, SIW, SKD, and UCG), participated in the experiment. All subjects were recruited by a notice on a bulletin board at Meisei University. Three females were undergraduate students and 1 was staff at the University. Three males were undergraduate students at Meisei University and 1 was a graduate. None of the subjects had any prior familiarity with the research topic. Before the experiment, the general procedures (settings, time period, payment, privacy, and the use of results) were explained to the subjects, who all signed a statement of informed consent.

Apparatus and Setting

Computer software made using Macromedia Director 8.0 controlled the stimulus presentations, procedural sequences, and response recording. The computer screen was gray and instructions were presented in black text unless otherwise noted below. An Apple Macintosh computer Power Macintosh G3/333 with a 17-inch monitor (TOTOKU CTD177OA-82A) was used. Speakers (Panasonic EAB-MPC70) were connected to the computer. The subjects used a pointing device, the Apple Desktop Mouse II, as a response indicator. The area in which the mouse was manipulated was covered with a box made of styrene foam so that the subjects could not see the movement of the mouse.

Subjects worked individually in a small room at Meisei University. The experimenter was in the next room, observing the subject via a one-way mirror. Three to five sessions, each lasting 20 to 60 min, were held per week. The total time needed to complete the experiment varied across subjects (see Table 1). When 30 min had elapsed in a session, the subject was given the option of terminating or completing up to an additional 30 min. Subjects earned 150 yen ($1.25) per 30 min of session time, plus 1.5 yen (1.2 cents) for every correct response. The cumulative amount earned (see Table 1) was not revealed to subjects until the experiment was completely finished.

Stimuli and Responses

Stimuli. The lower-right portion of Figure 1 shows the 12 symbols used as stimuli in the experiment. They appeared in a 7 cm by 7 cm square on the screen. For expository purposes, each stimulus in a class is assigned a letter (referring to the stimulus) and a number (referring to the potential equivalence class to which that stimulus could belong). Thus, for example, A1, B1, C1, and D1 were expected to become members of the same equivalence class.

Responses. Three response topographies were used, and are assigned a letter and a number for expository purposes (R1, R2, and R3). All responses began when the subject pressed the mouse button. The subject was then required to move the mouse in one of four directions (up, down, right, or left), while holding the button down. When the mouse moved approximately 1 cm in any of the four directions, a signal (440 Hz for 0.1 s) was presented to prompt the subject to change direction. The sequence of the direction was down-up-left-left for R1, left-right-up-down for R2, and right-down-right-up for R3. The responses ended when the mouse button was released.

Subjects were required to start a response after moving the on-screen cursor to a square in which a stimulus appeared. Once a response began, the cursor became invisible. If the cursor was moved out of the square during the response, the cursor became visible and the response was canceled. The subject then tried again. Virtual movement of the cursor was veridical to actual mouse movement. For example, when the subject moved the mouse 1 cm, the cursor moved approximately 1 cm on the screen.

Procedures

All instructions described were given in Japanese, but are shown below in English. The first session began with the following instructions displayed on the computer screen:

> Thank you for participating in the experiment. You will earn 1.5 yen whenever your response is correct. You will be informed of your total at the end of the experiment. If you have any questions, please ask the experimenter. Once the experiment begins, he will not answer questions. If you are ready, click on Go Next.

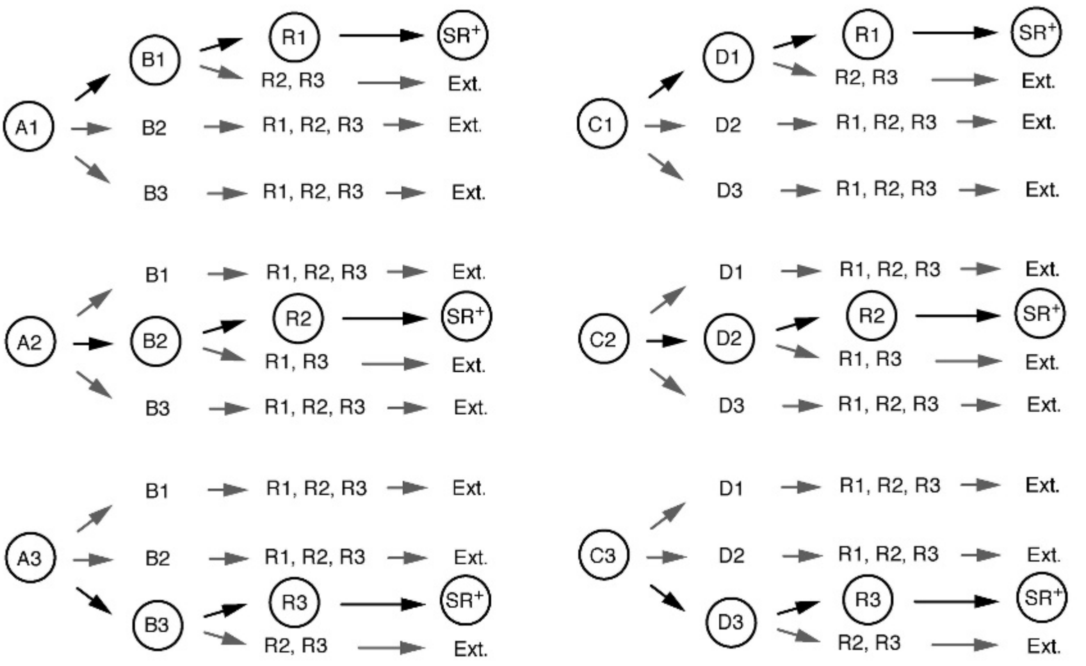
Response training. After initial instructions, the subjects were taught the three types of mouse movements described above, using procedures described in the Appendix.

Baseline training. A matching-to-sample procedure was used to establish the baseline relations among stimuli and responses. Two training sets were used, each represented by a series of 3 letters. For ABR training, stimuli A1, A2, and A3 were used as samples, stimuli B1, B2, and B3 were used as comparisons, and responses R1, R2, and R3 were used as differential responses. For CDR training, stimuli C1, C2, and C3 were used as samples, stimuli D1, D2, and D3 were used as comparisons, and responses R1, R2, and R3 were used as differential responses. At the beginning of baseline training, the following instructions were displayed:

> Whenever a square appears at the top of the screen, click on the square. Three squares will then appear at the bottom of the screen. Move the cursor to one of the three squares, and manipulate the mouse using one of the three ways you learned previously. If you want to retry the manipulation, move the cursor approximately 5 cm without releasing the button, and the cursor will appear out of the square. If you are ready, click on the Go Next.

At the beginning of each trial, a sample stimulus was presented at the top of the screen. Selection of the sample produced three comparison stimuli at the bottom of the screen. A correct response was defined as selecting a comparison stimulus in the same experimentally defined class as the sample, using the mouse movement topography designated as a part of that class. A correct response caused all of the stimuli to disappear, a chime to play briefly, and "Correct" to appear in red on the screen for 1.1 s. The next trial began after a 1.5-s intertrial interval. If the mouse movement or the selected comparison was incorrect, all of the stimuli disappeared,

FOUR-TERM CONTINGENCIES



EQUIVALENCE RELATIONS

THE EVENT PAIRS
(REFLEXIVE AND SYMMETRICAL RELATIONS)
(A1, A1), (B1, B1), (C1, C1), (D1, D1), (A2, A2), (B2, B2)
(C2, C2), (D2, D2), (A3, A3), (B3, B3), (C3, C3), (D3, D3)
(B1, A1), (B2, A2), (B3, A3), (D1, C1), (D2, C2), (D3, C3)

THE EVENT PAIRS INCLUDING RESPONSES
(R1, R1), (R2, R2), (R3, R3), (A1, R1), (B1, R1), (C1, R1)
(D1, R1), (A2, R2), (B2, R2), (C2, R2), (D2, R2), (A3, R3)
(B3, R3), (C3, R3), (D3, R3), (R1, A1), (R1, B1), (R1, C1)
(R1, D1), (R2, A2), (R2, B2), (R2, C2), (R2, D2), (R3, A3)
(R3, B3), (R3, C3), (R3, D3)

THE EVENT PAIRS COMBINED ACCORDING TO RESPONSES
(A1, C1), (A1, D1), (A2, C2), (A2, D2), (A3, C3), (A3, D3)
(C1, A1), (D1, A1), (C2, A2), (D2, A2), (C3, A3), (D3, A3)
(B1, C1), (B1, D1), (B2, C2), (B2, D2), (B3, C3), (B3, D3)
(C1, B1), (D1, B1), (C2, B2), (D2, B2), (C3, B3), (D3, B3)

SET OF STIMULI USED IN THIS STUDY

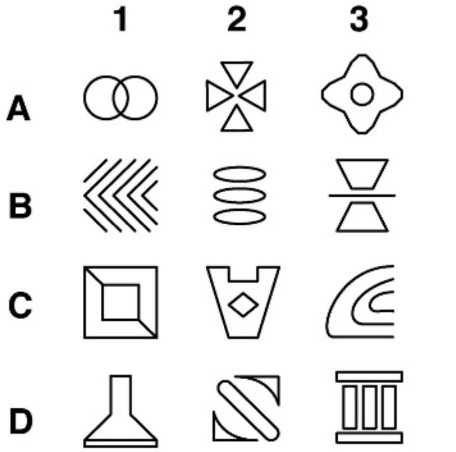


Table 1

Subject characteristics, duration of participation (number of sessions and total min), and total earnings in United States dollar equivalents.

Subject	Description, age	Sessions	Minutes	Earnings
SIW	Male undergraduate, 19	7	330	$45.92
SKD	Male undergraduate, 20	6	301	$41.29
SUT	Female undergraduate, 20	8	403	$46.12
UIR	Female undergraduate, 19	8	395	$47.49
IOH	Male postgraduate, 22	6	324	$42.53
MKB	Female University staff, 28	10	255	$42.92
TMM	Female undergraduate, 19	7	389	$52.43
UCG	Male undergraduate, 22	7	351	$46.52

a buzzer sounded, and ''Incorrect'' appeared in red for 0.3 s. The trial then repeated after the intertrial interval. Two consecutive correct responses on a correction trial were required to advance the session.

ABR training was conducted first, followed by CDR training. Both ABR and CDR training consisted of blocks of 18 trials. A brief tone and the Go Next button were presented as one block ended. Mastery criterion for both portions of training was 18/18 correct in one trial block. After ABR and CDR training, ABR and CDR trials were intermingled in blocks of 36 trials. Mastery criterion was 35/36 correct in one trial block. The computer presented ''Half finished'' for 5 s at the halfway point in each block during the intermixed condition.

The number of presentations of each sample stimulus was controlled for each block. Within a training block, the sample stimulus and the location of the correct comparison stimulus could remain constant for no more than three consecutive trials. In mixed training, the same relation (ABR or CDR) was not presented for more than three consecutive trials.

Following the mixed-training trial block in which the mastery criterion was met, the Go Next button was presented with a brief tone and the following message: ''The computer sometimes will not present feedback. If you are ready, click on the Go Next.'' The reinforcement probability was decreased in three steps (0.66, 0.33, and 0.00). Mastery criterion for each step was 35/36 correct in a single trial block. After each block was completed, the Go Next button and a brief jingle were presented.

One variation to this procedure was implemented when Subject TMM's response topographies became unstable when the reinforcement probability reached 0.00, causing her to fail to meet the mastery criterion across five trial blocks. Response training (Appendix) was reintroduced, followed by mixed training with reinforcement probability at 1.00. When the mastery criterion was met, reinforcement probability was reduced to 0.66 and then 0.00.

Equivalence tests. Equivalence tests consisted of 54 trials (18 trials of a potential emergent relation, and 18 trials each of the two trained relations, AB and CD), separated by a 1.5-s intertrial interval. No feedback was presented. On all trials, a subject could select any comparison using any response topography (e.g., it was possible to simply click once on a comparison to proceed to the next trial).

Prior to every test except for the initial BA test (see below), mixed training (ABR and CDR) with a 1.00 reinforcement probability was reintroduced, with 35/36 correct re-

←

Fig. 1. Four-term contingencies, equivalence relations, and sets of stimuli. A1, A2, A3, B1, B2, B3, C1, C2, C3, D1, D2, and D3 represent stimuli, and R1, R2, and R3 represent responses. SR+ indicates a reinforcer, and Ext. indicates extinction. Top: four-term contingencies in an arbitrary matching-to-sample procedure. A1, A2, A3, C1, C2, and C3 represent samples, and B1, B2, B3, D1, D2, and D3 represent comparisons. Black arrows illustrate units among which relations are strengthened by reinforcement; gray arrows illustrate units among which strengthening does not occur. Bottom, left: pairwise equivalence relations expected to emerge from four-term contingencies in the top panel. In the present study, 54 of these relations were tested and 15 relations (in which responses serve as comparison) were not tested. Bottom, right: sets of stimuli used in the present study. Numerals indicate membership in an equivalence class. Letters indicate how the stimuli were paired for introduction to the subjects.

Table 2

Number of trials in response training and baseline training.

Phase	Subjects							
	SIW	SKD	SUT	UIR	IOH	MKB	TMM	UCG
Response training								
Response 1	35	32	44	39	34	35	40	37
Response 2	32	31	32	37	30	39	35	30
Response 3	30	30	30	30	30	33	39	30
Practice	30	30	30	31	30	30	31	33
Baseline training								
ABR	236[a]	109[b]	67	79	118	85	268	125[c]
CDR	47	46	67	133	65	45	198	89
Random mix of ABR and CDR with reinforcement probability of 1.00	36	38	78	36	38	36	38	80
Random mix of ABR and CDR with decreasing reinforcement probability (0.66-0.33-0.00)	112	108	108	197	149	148	486[d]	110

[a] Not shown: SIW reviewed Response 1-2-3 training (Practice) for 30 trials between the 30th and 31st trials in ABR training. He also reviewed Response 1-2-3 training (Practice) for 33 trials between the 197th and 198th trials in ABR training.

[b] Not shown: SKD reviewed Response 1-2-3 training (Practice) for 30 trials between the 31st and 32nd trials in ABR training.

[c] Not shown: UCG reviewed Response 1-2-3 training (Practice) for 38 trials between the 35th and 36th trials in ABR training.

[d] Not shown: TMM reviewed Response 1-2-3 training (Practice) for 31 trials, and mixed training with a reinforcement probability of 1.00 for 78 trials, between the 374th and 375th trials in the mixed training with a decreasing reinforcement probability.

sponses required to proceed to the next test. After the 27th trial of each test, the computer presented the words ''Half finished'' for 5 s. Upon completion of a test, a tone played briefly and the Go Next button appeared at the bottom of the screen; clicking on this button advanced the session to the next test.

The first set of tests involved symmetrical relations. The first test focused on BA trials, with stimuli B1, B2, and B3 presented as samples on different trials and stimuli A1, A2, and A3 presented as comparisons on all trials. The second test focused on DC trials, with stimuli D1, D2, and D3 presented as samples on different trials and stimuli C1, C2, and C3 presented as comparisons on all trials.

The second set of tests involved relations that included differential responses as a sample (RA, RB, RC, and RD). Prior to the first trial, the following instructions were presented:

> At times you will see a white square not filled with a drawing at the top of the screen. Whenever you see this square, move the cursor to the square and manipulate the mouse in one of the three ways you learned previously. If your manipulation is correct, three squares will appear at the bottom of the screen. Respond to one of the three squares as you did before. If you are ready, click on the Go Next.

At the beginning of each trial, a white square was presented at the top of the screen. Moving the cursor to the square and employing the response topography experimentally defined as correct on that trial produced the comparison stimuli. Other mouse movements were ineffective in advancing the trial.

The third set of tests involved relations in which the stimuli were associated with a common response topography (AC, CA, AD, DA, BC, CB, BD, and DB). The fourth set of tests involved reflexive relations (AA, BB, CC, and DD). Trials in these two test sets were structured similarly to those in the first set of tests. After all of the tests were completed, a message was displayed indicating that the experiment had concluded and thanking the subject for participating.

RESULTS

Response Training and Baseline Training

Table 2 shows the number of trials required to meet the mastery criterion in response

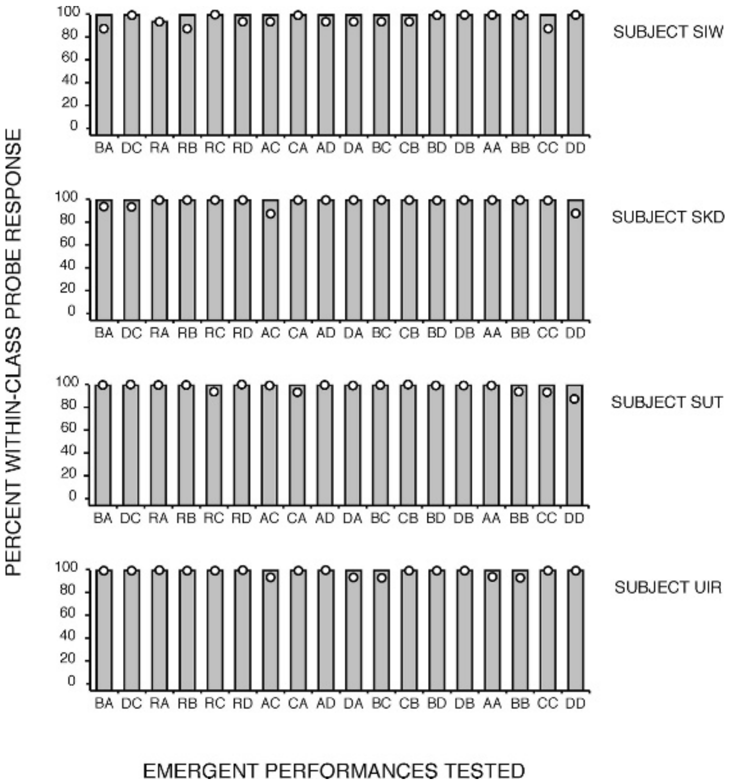


Fig. 2. Percentages of responses in test trials that were consistent with equivalence classes A1–B1–C1–D1–R1, A2–B2–C2–D2–R2, and A3–B3–C3–D3–R3 for subjects who showed no emergence failures. Bars show the results of 18 probe trials for the sample–comparison relation. Open circles show the results for the sample–comparison–response relation.

training and baseline training. If a subject committed no errors during response training, the total number of trials was 30 (Appendix), and most subjects met the mastery criterion with few errors (median number of trials = 31.5). During baseline training, 4 of 8 subjects (SUT, UIR, IOH, and MKB) progressed without additional training, whereas a review of response training was required for the other four subjects.

Test

Maintenance of baseline relations was assessed along with emergence of untrained relations. Baseline relations were considered to have remained intact if demonstrated at a level of at least 30/36 (83%) correct in a test. This criterion was met by all subjects in all test blocks.

Two results were of interest during the equivalence tests: first, the extent to which subjects selected comparison stimuli (using any response topography) indicating the emergence of untrained sample–comparison relations; and second, the extent to which subjects used class-consistent response topographies to select comparison stimuli, suggesting the emergence of untrained sample–comparison–response relations.

Bars in Figures 2 and 3 show the percentages of comparison selections that were consistent with untrained sample-comparison relations in equivalence classes A1–B1–C1–D1–R1, A2–B2–C2–D2–R2, and A3–B3–C3–D3–R3. The order of bars from left to right indicates the order in which the tests were conducted. Each bar represents the results of 18 test trials, six for each of three relations. Thus, for example, the BA results incorporate tests of B1–A1, B2–A2, and B3–A3. A subject was considered to have passed a test—to have demonstrated emergence of an untrained relation—if the result for all three component relations was at least 5/6 correct. Otherwise,

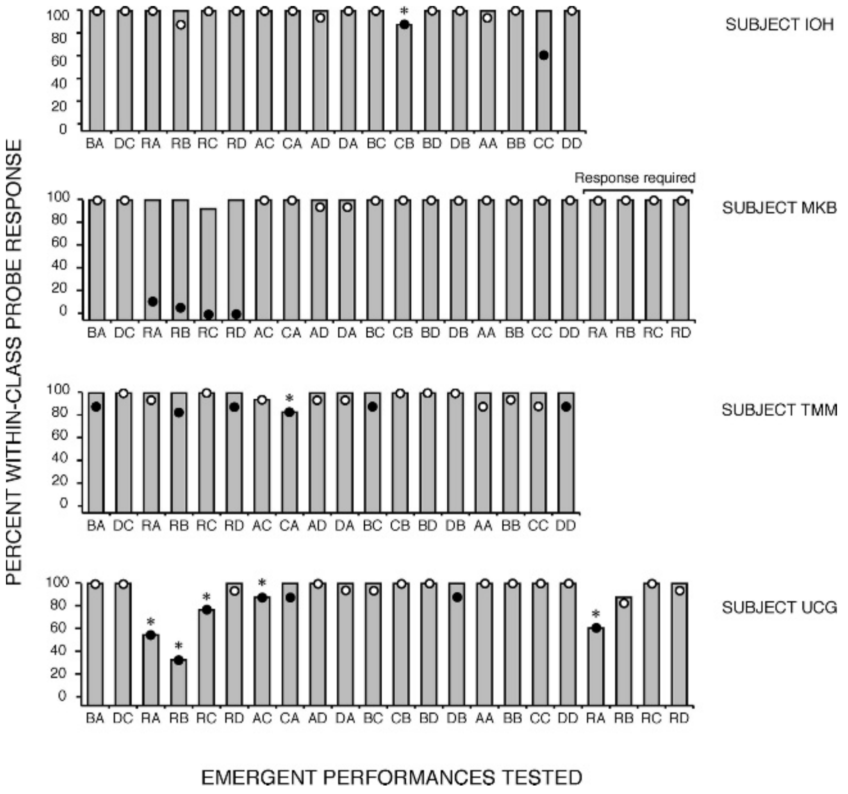


Fig. 3. Percentages of responses in test trials that were consistent with equivalence classes A1–B1–C1–D1–R1, A2–B2–C2–D2–R2, and A3–B3–C3–D3–R3 for subjects who showed some emergence failures. Bars show the results of 18 probe trials for the sample–comparison relation. Asterisks at the top of the bars indicate that the result of the sample–comparison relation was negative. Circles show the results of the sample–comparison–response relation. Open circles indicate that the result was positive, whereas filled circles indicate that the results were negative.

the outcome was characterized as an emergence failure. In the figures, an asterisk above a bar indicates an emergence failure.

Figures 2 and 3 (circles) also show the results of tests for sample–comparison–response relations. Recall that differential response topography was required to generate sample stimuli on each trial. Thereafter, any response topography could be used to select a comparison stimulus. A sample–comparison–response relationship was said to have emerged if this selection was made using the class-consistent differential response topography—for example, in the BA test, if R1 was used to select comparison A1 conditionally upon sample B1. An unfilled circle above a bar indicates successful emergence of a sample–comparison–response relation (at least 5/6 class-consistent responses), whereas a filled circle indicates an emergence failure.

Figure 2 shows that subjects SIW, SKD, SUT, and UIR demonstrated all possible untrained relations, including sample–comparison relations with stimuli as samples, sample–comparison relations with responses as samples, and sample–comparison–response relations.

Emergence failures. Figure 3 shows the results for subjects IOH, MKB, TMM, and UCG, all of whom showed some emergence failures. MKB demonstrated emergence of all sample–comparison relations (bars), whereas IOH and TMM demonstrated emergence of all but one of these relations. UCG showed emergence failure for 4/18 sample-comparison relations. Sample–comparison emergence failures were as follows. IOH scored 16/18 (88%) correct on the CB test but failed to meet the mastery criterion for the C2–B2 relation (4/6). TMM scored 15/18 (83%) correct on the CA test but failed to meet criterion for the C1–A1 relation (3/6). UCG scored 16/18 (88%) correct on the AC test but failed to meet criterion for the A3–C3 relation (4/6). Additionally, subject UCG was unique in failing three of the four sample-comparison relations in which responses served as samples (RA, RB, and RC), although two of the three untrained relations emerged upon retesting.

All subjects in Figure 3 showed emergence of at least 12, and as many as 16, of the 18 untrained sample–comparison–response relations (circles). IOH showed emergence failures for two relations involving visual samples and comparisons. For TMM, MKB, and UCG, emergence failures occurred mainly for relations in which responses served as samples. Following completion of the test battery, tests of these relations (RA, RB, RC, and RD) were repeated for UCG and MKB. Due to experimenter oversight, the same tests were not repeated for TMM.

Inspection of UCG's initial test results suggested a bias for selecting the center comparison stimulus during early trials involving this relation type (results not shown), but this tendency appeared to deteriorate across trials. Thus, UCG's second test of sample–comparison–response relations was identical to the first. Upon retesting, this subject showed emergence of two of the three relations that previously had failed to emerge.

During the initial tests, MKB had selected comparison stimuli on the RA, RB, RC, and RD relations using a single, undifferentiated mouse click, prompting a modification of the procedure for the replication. The computer now required MKB to perform one of the three differential responses (R1, R2 or R3) when a comparison was selected. Upon retesting, MKB demonstrated emergence of all sample–comparison–response relations involving responses as samples.

DISCUSSION

The present study examined response–stimulus relations using an experimental design modeled after one proposed by Sidman (2000). A major technical challenge in conducting such an experiment lies in arranging for responses to serve as sample stimuli. The present study met this challenge by employing a variant of an unsignaled differential-response procedure used previously with pigeons by Lionello-DeNolf and Urcuioli (2003). The results showed reliable emergence of symmetry and reflexivity relations, as well as, in most subjects, response–stimulus relations.

Sidman (2000) proposed that an equivalence relation may include ordered pairs of all positive components of a reinforcement contingency, including a stimulus, a response, and a reinforcer. Most published studies have examined relations only among stimuli, but a few reports suggest that reinforcers also may be members of equivalence classes (Dube & McIlvane 1995; Dube et al., 1987; Dube et al., 1989; Goyos 2000; McIlvane et al., 1992; Schenk 1994). The present study provides evidence that an equivalence relation, as defined by Sidman (1971), also may include nonverbal responses in humans. Response–stimulus relations emerged in 7 of 8 subjects in the RA, RB, RC, and RD tests, in which differential responses were used as a sample stimulus. The subject who initially failed to show response–stimulus relations then did so when the relevant tests were repeated. Moreover, in most cases, emergent relations were demonstrated in which the sample and comparison stimuli were related only via association with a common response topography (AC, CA, AD, DA, BC, CB, BD, and DB tests). In 64 relevant tests conducted across 8 subjects, emergence was demonstrated in 61 cases.

One requirement for tests of emergent stimulus–stimulus relations is that the tested relations have no direct reinforcement history. The same rule can be extended to tests for emergent relations including responses. Manabe et al. (1995), for example, reported that a sample–response–comparison relation emerged in budgerigars. The budgerigars made low- or high-frequency calls in response to sample stimuli for which those calls had not been reinforced previously. However, one cannot exclude the possibility that the response–comparison relations may have been established prior to the test because the same response–comparison relations had been reinforced under a different sample before the test (Saunders & Williams, 1998). In contrast, the emergent relations in the present study had no reinforcement history prior to the tests, and thus the results provide strong evidence that equivalence relations may include a response.

Some results suggest that the members of an equivalence class are functionally interchangeable (e.g., Sidman et al., 1985), and the present results may be interpreted as supporting this proposition. All subjects demonstrated differential responses during the tests when they selected a comparison using class-consistent response topographies—for example, in the AC test, when selecting comparison C1 conditionally upon sample A1 with response R1. In baseline training, the stimuli (B1, B2, B3, and D1, D2, D3) served as discriminative stimuli for responses R1, R2 and R3, respectively. The sample stimuli (A1, A2, A3, and C1, C2, C3) served as conditional discriminative stimuli controlling comparison selection. In tests, however, stimuli that previously served as samples also controlled differential response topographies and served as comparison stimuli.

One subject, TMM, showed unreliable emergence of sample–comparison–response relations (Figure 3), and this may have been due at least partly to poor mouse-control skills. The computer detected, and recorded, mouse movement when the virtual position of the cursor changed by 10 pixels (about 1 cm on the screen). Post hoc examination of session records from the equivalence tests showed that TMM often approximated the experimentally defined response topographies in ways that were counted as incorrect by the computer. For example, when she performed R1 (''down-up-left-left''), in some trials the computer detected ''down-up-left-left-left'' because of extra mouse movement during the final link in the response chain. Similarly, when she performed R3 (right-down-right-up), the computer sometimes detected right-down-down-right-up, because of extra movement in the second link. Acquisition of mouse-control skills rarely is addressed in reports of computer-based human-operant research, although adverse effects of poor mouse-control skills on human operant performance have been described previously (e.g., Ecott & Critchfield, 2004). Experimenters appear to assume that their typically developing adult subjects will possess adequate mouse-control skills, but TMM's difficulties, which arose despite the completion of a fairly rigorous response training procedure (Appendix), demonstrate that this is not always the case. Had TMM been given additional training for mouse control, or simply been forced to execute one of the three experimentally defined movement sequences to advance each trial during the tests (as was the case in MKB's test replications), emergence of at least some sample–comparison–response relations might have been demonstrated.

When a response serves as a discriminative stimulus or a conditional discriminative stimulus in an experiment, what is the real antecedent event? Presumably, response-produced stimuli are implicated, but several mundane possibilities can be ruled out. For example, auditory feedback followed each mouse movement, but this feedback was identical for all mouse movements, regardless of response topography, and therefore could not have served as a conditional discriminative stimulus. Additionally, a procedure like that of the present study might include response-associated cursor movements displayed on the subject's screen, but in the present study, the cursor disappeared while the subject performed a differential response. Finally, a subject might observe his or her own hand performing the mouse manipulation, but the mouse hand was shielded from view while the differential response topographies were executed. Thus, auditory feedback and visual detection of cursor and hand movements can be ruled out as conditional discriminative stimuli. Presumably, interoceptive feedback (see Skinner, 1974) integral to performing the mouse movements served as the conditional discriminative stimuli.

Regardless of which aspect of the response served as the conditional discriminative stimulus, the present study joins many others in showing that one instance of behavior can exert influence over another. For example, in a study by Shimp (1983), pigeons pecked at a key with either a long or short interresponse interval (IRI). Following this response, two comparison stimuli were presented. Selection of one was reinforced if the preceding IRI had been long, and selection of the other was reinforced if the IRI had been short. The results were discussed in terms of self-reporting. In a broadly similar study, the results were discussed in terms of memory (Fetterman & MacEwen, 1989). In both cases, however, responses can be considered more economically as conditional discriminative stimuli (e.g., see Branch, 1977; Critchfield & Perone, 1990).

Conditional discriminative control also clearly operates in studies of response variability, in which a key determinant of variability is the reinforcement schedule, often including a contingency in which responses are reinforced only when the topography differs from that of several preceding responses (Page & Neuringer, 1985). Moreover, the degree of variability can be brought under control of a discriminative stimulus (Denney & Neuringer, 1998). These outcomes suggest that response variability might serve as either sample or comparison in conditional discrimination training required to establish equivalence classes, thereby increasing the complexity of behavior that participates in equivalence classes.

In summary, stimuli, reinforcers, and responses now have all been implicated in equivalence classes, but it is important to note that not all possible relations have yet been adequately tested. For example, in the present experiment, several potential emergent equivalence relations involving responses as comparison stimuli were not tested because doing so is procedurally challenging (see Dube et al., 1993). Additionally, it is not known whether equivalence relations can be established between reinforcers and responses. Much experimental work remains in the quest to fully test Sidman's (2000) proposal that stimuli, responses, and reinforcers all are integrated in equivalence classes.

REFERENCES

Branch, M. N. (1977). On the role of ''memory'' in the analysis of behavior. *Journal of the Experimental Analysis of Behavior, 28*, 171–179.

Critchfield, T. S., & Perone, M. (1990). Verbal self-reports of delayed matching to sample by humans. *Journal of the Experimental Analysis of Behavior, 53*, 321–344.

Denney, J., & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior, 26*, 154–162.

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 59*, 103–114.

Dube, W. V., & McIlvane, W. J. (1995). Stimulus-reinforcer relations and emergent matching to sample. *The Psychological Record, 45*, 591–612.

Dube, W. V., McIlvane, W. J., Mackay, H. A., & Stoddard, L. T. (1987). Stimulus class membership established via stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 47*, 159–175.

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., & Stoddard, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 51*, 65–76.

Ecott, C. L., & Critchfield, T. S. (2004). Noncontingent reinforcement, alternative reinforcement, and the matching law: A laboratory demonstration. *Journal of Applied Behavior Analysis, 37*, 249–265.

Eikeseth, S., & Smith, T. (1992). The development of functional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of the Experimental Analysis of Behavior, 58*, 123–133.

Fetterman, J. G., & MacEwen, D. (1989). Short-term memory for responses: The ''choose-small'' effect. *Journal of the Experimental Analysis of Behavior, 52*, 311–324.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48*, 317–332.

Goyos, C. (2000). Equivalence class formation via common reinforcers among preschool children. *The Psychological Record, 50*, 629–654.

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 51*, 385–392.

Lionello-DeNolf, K. M., & Urcuioli, P. J. (2003). A procedure for generating differential ''sample'' responding without different exteroceptive stimuli. *Journal of the Experimental Analysis of Behavior, 79*, 21–35.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior, 78*, 527–549.

Manabe, K., Kawashima, T., & Staddon, J. E. R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior, 63*, 111–126.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., deRose, J. C., & Stoddard, L. T. (1992). Stimulus-reinforcer relations and conditional discrimination. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 43–67). Reno, NV: Context Press.

McIntire, K. D., Cleary, J., & Thompson, T. (1987). Conditional relations by monkeys: Reflexivity, symmetry, and transitivity. *Journal of the Experimental Analysis of Behavior, 47*, 279–285.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 429–452.

Saunders, K. J., & Williams, D. C. (1998). Do parakeets exhibit derived stimulus control? Some thoughts on experimental control procedures. *Journal of the Experimental Analysis of Behavior, 70*, 321–324.

Schenk, J. J. (1994). Emergent relations of equivalence generated by outcome-specific consequences in conditional discrimination. *The Psychological Record, 44*, 537–558.

Shimp, C. P. (1983). The local organization of behavior: Dissociations between a pigeon's behavior and self-reports of that behavior. *Journal of the Experimental Analysis of Behavior, 39*, 61–68.

Sidman, M. (1971). Reading and auditory-visual equivalence. *Journal of Speech and Hearing Research, 14*, 5–13.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213–245). Mahwah, NJ: Lawrence Erlbaum Associates.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74,* 127–146.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43,* 21–42.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5–22.

Skinner, B. F. (1974). *About behaviorism.* New York: Vintage Books.

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56,* 519–555.

Received: January 18, 2003
Final acceptance: June 19, 2006

APPENDIX

Initial Training

Three response topographies were trained, always in the order of R1, then R2, then R3. Each topography was trained in three stages.

Stage 1. At the beginning of response training, the following message was displayed at the top of the screen:

> You will learn mouse manipulation. Instructions will appear on the screen at the beginning, but won't appear later. Try to manipulate the mouse without the instructions. If you are ready, click on Go Next.

In the first stage, mouse movement was prompted by printed instructions and black arrows. Instructions that appeared at the bottom of the screen were in black text, whereas those that appeared at the top of the screen were in red text. A trial began with the presentation of a white square (7 cm × 7 cm) at the center of the screen, and instructions at the bottom of the screen stating, ''Move the cursor to the square, without pressing the mouse button''. Doing so changed the message at the bottom of the screen to, ''Set the cursor at the center of the square, press the mouse button, and keep holding the button''. Pressing the button removed the cursor from the screen and produced an arrow accompanied by the instruction, ''While holding the button, move the mouse in the direction of the arrow until you hear a sound''. When the mouse was moved about 1 cm, a 0.1-s tone (440 Hz) sounded. If the direction of the mouse movement matched that indicated by the arrow, the arrow changed direction to indicate the direction of the next required movement. The subject was required to change direction four times. If the subject correctly changed direction four times, the instruction ''Release the button'' appeared. When the subject released the button, the computer presented a brief chime and the word ''correct'' appeared in red at the center of the screen for 1.1 s. After a 1.5-s ITI, the next trial began.

During the first stage, several feedback messages promoted correct mouse movements. If the mouse was moved in an incorrect direction, the arrow disappeared and a message at the top of the screen stated, ''Your manipulation is incorrect. Move the mouse approximately 5 cm, while holding the button until the cursor appears outside the square, and retry the manipulation''. Although the cursor could not be seen during mouse movement, the computer was constantly monitoring its virtual position. If the cursor was moved out of the square, it became visible and a message at the top of the screen stated, ''The cursor has moved out of the square. Make sure that the cursor is not in the square, release the button, and then retry the manipulation.'' With these instructions, two more instructions were presented at the bottom of the screen. If the cursor was visible and the subject moved the cursor to the square while holding the button, then the following message appeared: ''Move the cursor out of the square, and release the button.'' If the cursor was out of the square and the button was depressed, then the following message appeared: ''Move the cursor to the square, without holding the mouse button.'' Following any of these instructions, the subject could retry the manipulation.

A mouse movement that resulted in a retry was not counted as an incorrect response. Once a movement was initiated, the cursor became invisible and the mouse button had to remain depressed. Release of the button meant the end of the movement. If the button was released during mouse movement or after an incorrect movement, an incorrect response was registered, a brief buzzer sounded, and ''Incorrect'' appeared in red in the center of the screen for 0.3 s. After a 1.5-s ITI, the next trial began.

Stage 2. At the start of the second stage, a 5-s printed message stated, ''The presentation of the instructions will be delayed. Don't wait for the instructions and manipulate the mouse as before.'' Consistent with the instructions, during this stage, the presentation of printed feedback and instructions was delayed for 3 s following a response, with one exception: Following incorrect mouse movements or movements of the virtual cursor out of the square, feedback messages identical to those of Stage 1 appeared immediately at the top of the screen, and the trial was counted as incorrect.

Stage 3. The following message was displayed in black for 5 s before the third stage: ''Instructions will not appear.'' Consistent with this message, during the third stage no instructions were given. The only feedback provided by the computer for mouse movement was the words ''correct'' or ''incorrect.''

Mastery criterion. Training began with a counter on the subject's screen stating, ''30 trials left.'' Each correct response decremented the counter by one. An incorrect response incremented the counter to the next higher multiple of 5 (e.g., an incorrect response with the counter set at 17 would raise the counter to 20). Training continued until the counter reached zero.

Practice

Once the third topography had been trained, a practice period was introduced with the following message: ''You will now practice the three mouse manipulations that you have learned. If you are ready, click on the Go Next.'' All subjects practiced R1, then R2, then R3. Each response was trained with two stages. During the first stage, the feedback messages described above were presented immediately following a response. The stage continued until five correct responses had been made. During the second stage, no instructions were presented. The stage continued until five consecutive correct responses had been made.

Additional Prompts and Training

At any point during the experiment, if the mouse was placed diagonally in its tray, the computer could not correctly detect mouse movement. When this happened, the experimenter stated: ''Don't place the mouse diagonally.'' This occurred once each for IOH, SKD, SUT, TMM, and UIR during response training or baseline training.

During response training, subjects occasionally performed erroneous mouse movements after apparently not reading the instructions on the screen. In these cases, instructions describing how to retry the mouse movements were presented in writing (twice for SIW, TMM, and UCG, and once for SKD) or verbally (twice for MKB). The instructions stated, ''Your manipulation is incorrect. Move the mouse approximately 5 cm, while pressing the button until the cursor appears out of the square. Make sure that the cursor is not in the square, release the button, and retry the manipulation.''

During the ABR baseline training that followed completion of the response training, a review of response training was instituted for 3 subjects who employed incorrect response topographies on at least 30 trials. This occurred once for SKD and UCG and twice for SIW. Additional response training also was provided for TMM after she failed to meet the mastery criterion for five consecutive trial blocks during the mixed training (reinforcement probability = 0.00) that preceded emergent-relations tests.